Response to questions

Brattin, Bill to: Kopylev.Leonid, Benson.Bob 10/02/2012 10:31 AM

From: "Brattin, Bill" <brattin@srcinc.com>
To:
Cc: DeVoney.Danielle@epamail.epa.gov, Berry.David@epamail.epa.gov, Christensen.Krista@epamail.epa.gov, Bateson.Thomas@epamail.epa.gov

---

Leonid
Here are my answers to your questions/comments:

<![if !supportLists]>1)<![endif]>Linear function doesn't need to go below zero. It could be max(0, linear function) or max (min IH data, linear function) and thus not disqualified on being below 0.
Sure, we can substitute zero for negative values. Better (I think) is to constrain the linear fit so the predicted values never go negative. This is a little tricky since what is being constrained in the output, not the parameter per se. I have tried that manually, and it really does not change much.

2) I think fitting linear function to outdoor was discussed, was it? Was GOF evaluated formally?
I will test a 2-segment linear on the outdoor jobs and send the results for consideration
Question: what GOF statistic will be useful? The only one that I can compute easily is mean square error. Is that sufficient?

3) I am not sure how Fig. 2 was generated for jobs other than expander - could more details be provided?
Figure 2 shows the results of a simultaneous fit of 3-segment exponential to all 7 indoor jobs, and also a simultaneous fit of a 2-segment exponential model to the two outdoor jobs.
I did this in Excel using Solver. Conceptually, for the indoor jobs, we are fitting 3 curves (2 parameters each) to each of the 7 data sets = 42 parameters. However, for each job, the 3 curves are constrained so that they intersect at the junction between the time segments. This reduces the number of fitting parameters from 42 to 10 (there are 7 intercept terms (one for each job) and 3 slope terms (one for each segment). I have attached a Word document that I prepared for discussion with UC that tries to explain this. It seems a little complicated, but is actually quite simple and the fitting works easily. If it would be useful, I can send the Excel sheet with the fitting in it, but you have to know Excel and be familiar with Solver for this to be very helpful.

4) Is it correct to assume that fit was done in such a way as fit to data to 1980 could be evaluated by dropping the last curve or is it more complicated than that? I.e. how continuity was achieved?
See response above. The fitting is simultaneous across all 3 segments, including post 1980 data. NOTE: if the exponential parameters are unconstrained, the slope for the post 1980 data (-b3) is very slightly positive (increasing concentrating with time). To avoid that, I constrained all of the slope terms so the b terms must by >= 0.001 (3rd segment slope = -.001). The slope is so shallow here that we could actually just treat C(post 1980) as a constant.

****

Bill Brattin
SRC, Inc.

999 18th Street Suite 1150
Denver CO 80202
Phone: 303-357-3121
Fax: 303-292-4755
e-mail: brattin@srcinc.com
From: Leonid Kopylev [mailto:Kopylev.Leonid@epamail.epa.gov]
Sent: Monday, October 01, 2012 1:53 PM
To: Bob Benson
Cc: Brattin, Bill; Danielle DeVoney; David Berry; HILBERTJ@UCMAIL.UC.EDU; Krista Christensen; Thomas Bateson
Subject: Re: Recommended AM based JEM
Bob and Bill:

thanks for the revised text. Few quick questions (to add to Krista's)
1) Linear function doesn't need to go below zero. It could be max(0, linear function) or max (min IH data, linear function) and thus not disqualified on being below 0.
2) I think fitting linear function to outdoor was discussed, was it? Was GOF evaluated formally?
3) I am not sure how Fig. 2 was generated for jobs other than expander - could more details be provided?
4) Is it correct to assume that fit was done in such a way as fit to data to 1980 could be evaluated by dropping the last curve or is it more complicated than that? I.e. how continuity was achieved?

Thanks,
Leonid

From: Bob Benson/R8/USEPA/US
To: Thomas Bateson/DC/USEPA/US@EPA, Krista Christensen/DC/USEPA/US@EPA, Leonid Kopylev/DC/USEPA/US@EPA, Danielle DeVoney/DC/USEPA/US@EPA, HILBERTJ@UCMAIL.UC.EDU
Cc: David Berry/R8/USEPA/US@EPA, brattin@srcinc.com
Date: 10/01/2012 12:58 PM
Subject: Recommended AM based JEM

Thank you all for the productive discussion last Thursday! We have consolidated the discussion points into a full proposal for developing the JEM based on the arithmetic mean of the IH data sets. The recommended approach and data plots are attached. We will try to answer any questions you have.

We think this recommended approach has the following advantages:
1) It is qualitatively similar to the approach used by UC in deriving the GM-based JEM
2) It uses the IH data in a scientifically defensible manner
3) It uses the information on engineering controls put in place at various dates

in a defensible manner
4) The plots show the fits are reasonable

We do not know if Linda can implement the variance weighted calculation in SAS. If that is possible, we will use the variance weighted calculations. If not, we will use the un-weighted calculations that we have now.

We do not believe that additional discussion will reveal a superior approach. Therefore, we are asking for your concurrence with the recommended approach by COB October 5 or before.

[attachment "Proposed Approach for JEM Oct 2012 v2.doc" deleted by Leonid Kopylev/DC/USEPA/US]